Ribosomal protein S7: a new RNA-binding motif with structural similarities to a DNA architectural factor

Harumi Hosaka[1], Atsushi Nakagawa[1], Isao Tanaka[1]*, Nao Harada[2], Kazunari Sano[2], Makoto Kimura[2], Min Yao[3,4] and Soichi Wakatsuki[3]

Background: The ribosome is a ribonucleoprotein complex which performs the crucial function of protein biosynthesis. Its role is to decode mRNAs within the cell and to synthesize the corresponding proteins. Ribosomal protein S7 is located at the head of the small (30S) subunit of the ribosome and faces into the decoding centre. S7 is one of the primary 16S rRNA-binding proteins responsible for initiating the assembly of the head of the 30S subunit. In addition, S7 has been shown to be the major protein component to cross-link with tRNA molecules bound at both the aminoacyl-tRNA (A) and peptidyl-tRNA (P) sites of the ribosome. The ribosomal protein S7 clearly plays an important role in ribosome function. It was hoped that an atomic-resolution structure of this protein would aid our understanding of ribosomal mechanisms.

Results: The structure of ribosomal protein S7 from *Bacillus stearothermophilus* has been solved at 2.5 Å resolution using multiwavelength anomalous diffraction and selenomethionyl-substituted proteins. The molecule consists of a helical hydrophobic core domain and a β-ribbon arm extending from the hydrophobic core. The helical core domain is composed of a pair of entangled helix-turn-helix motifs; the fold of the core is similar to that of a DNA architectural factor. Highly conserved basic and aromatic residues are clustered on one face of the S7 molecule and create a 16S rRNA contact surface.

Conclusions: The molecular structure of S7, together with the results of previous cross-linking experiments, suggest how this ribosomal protein binds to the 3′ major domain of 16S rRNA and mediates the folding of 16S rRNA to create the ribosome decoding centre.

Addresses: [1]Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060, Japan, [2]Laboratory of Biochemistry, Faculty of Agriculture, Kyushu University, Fukuoka 812, Japan, [3]European Synchrotron Radiation Facility, BP 220, F-38043, Grenoble Cedex, France and [4]MAC Science, 1-5-1 Shinyokohama, Kohoku-ku, Yokohama, Kanagawa 222, Japan.

*Corresponding author.
E-mail: tanaka@polymer.hokudai.ac.jp

Key words: decoding center, ribosomal protein S7, RNA architectural factor, RNA-binding protein, X-ray structure

Received: **30 June 1997**
Revisions requested: **31 July 1997**
Revisions received: **11 August 1997**
Accepted: **11 August 1997**

Structure 15 September 1997, **5**:1199–1208
http://biomednet.com/elecref/0969212600501199

Introduction

The ribosome is a ribonucleoprotein complex responsible for protein synthesis in all living cells. Its fundamental role is to decode mRNAs and synthesize proteins according to their message. The ribosome comprises two subunits, the large (50S) and small (30S) subunits. The functional region responsible for decoding and peptidyltransferase activity is located between the two subunits. An immunoelectron microscopic study showed that the decoding site is positioned deep in the cleft between the head and the lateral protrusion of the small subunit [1]. Much better three-dimensional images of the whole ribosome particle have been obtained recently by cryo-electron microscopy (cryo-EM) [2,3]. The orientations of tRNA molecules bound at the aminoacyl-tRNA (A), peptidyl-tRNA (P) and exit (E) sites of the ribosome have also been discussed [4,5]. To elucidate the detailed mechanisms of ribosome function, however, an atomic-resolution structure of the functional region is necessary. As a step towards trying to construct a detailed three-dimensional structure of the ribosome, biochemical methods such as cross-linking and footprinting have been extensively used [6]. These studies were used to establish the positions of ribosomal proteins, which comprise one third of the total mass of the ribosome, in relation to 16S rRNA, however, they often did not identify precise amino acid locations. Atomic-resolution structures of ribosomal proteins are thus indispensable for understanding the mechanism of ribosome function.

Ribosomal protein S7 is located on the side of the head of the 30S subunit facing the lateral protrusion, just above the cleft and the decoding site [7–10]. It is one of two protein components which are responsible for initiating the assembly of the 30S subunit; its binding to the 16S rRNA is crucial for the assembly of the head of the 30S subunit [11,12]. S7 is also a major protein component of the ribosome which has been shown to cross-link with tRNA molecules bound at the A and P sites [13–16]. A cross-linking experiment using mRNA analogues, in which the photoreactive label is located on the 5′-side of the coding triplets, shows that S7 is the only notable target which is cross-linked to mRNA [17]. S7 is also one of the principal regulatory elements that controls ribosomal protein synthesis by the translational feedback mechanism

[18]: it regulates the expression of the *str* operon that encodes ribosomal proteins S12 and S7, as well as elongation factors G and Tu. S7 binds to mRNA at the S12–S7 intercistronic region and the nucleotide sequence of this mRNA binding site has limited similarities to the S7-binding site on the 16S rRNA [19].

In this paper we present for the first time the atomic-resolution structure of a protein which is involved directly in the decoding centre of the ribosome. The structure of ribosomal protein S7 from *Bacillus stearothermophilus* was determined by the multiwavelength anomalous diffraction (MAD) method using selenomethionyl substituted protein. The tertiary structure of S7 contains features which may explain how it binds to 16S rRNA and also how it initiates folding of this RNA.

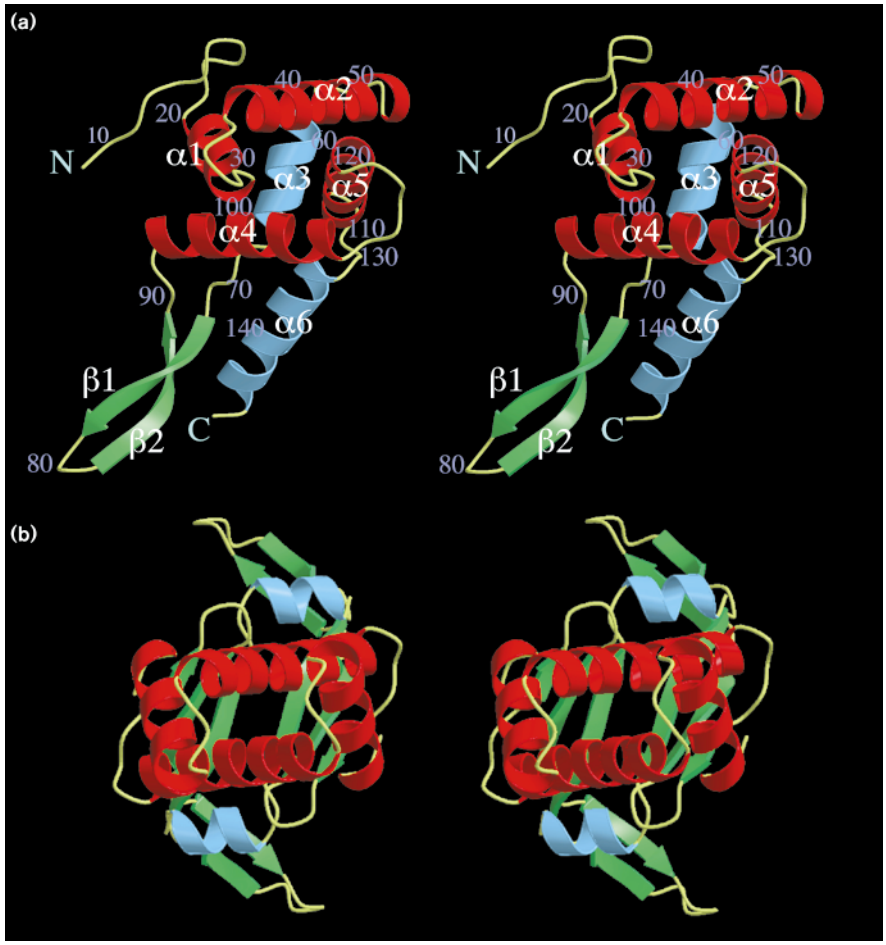
Results and discussion

Structure description

A stereo view of the overall structure of the S7 protein is given in Figure 1. The molecule consists of an α-helical domain with a β ribbon extending from it. The overall shape of the molecule is elliptical with dimensions of 50 Å along the long axis and 30 Å × 25 Å along the other two axes of the α-helical domain.

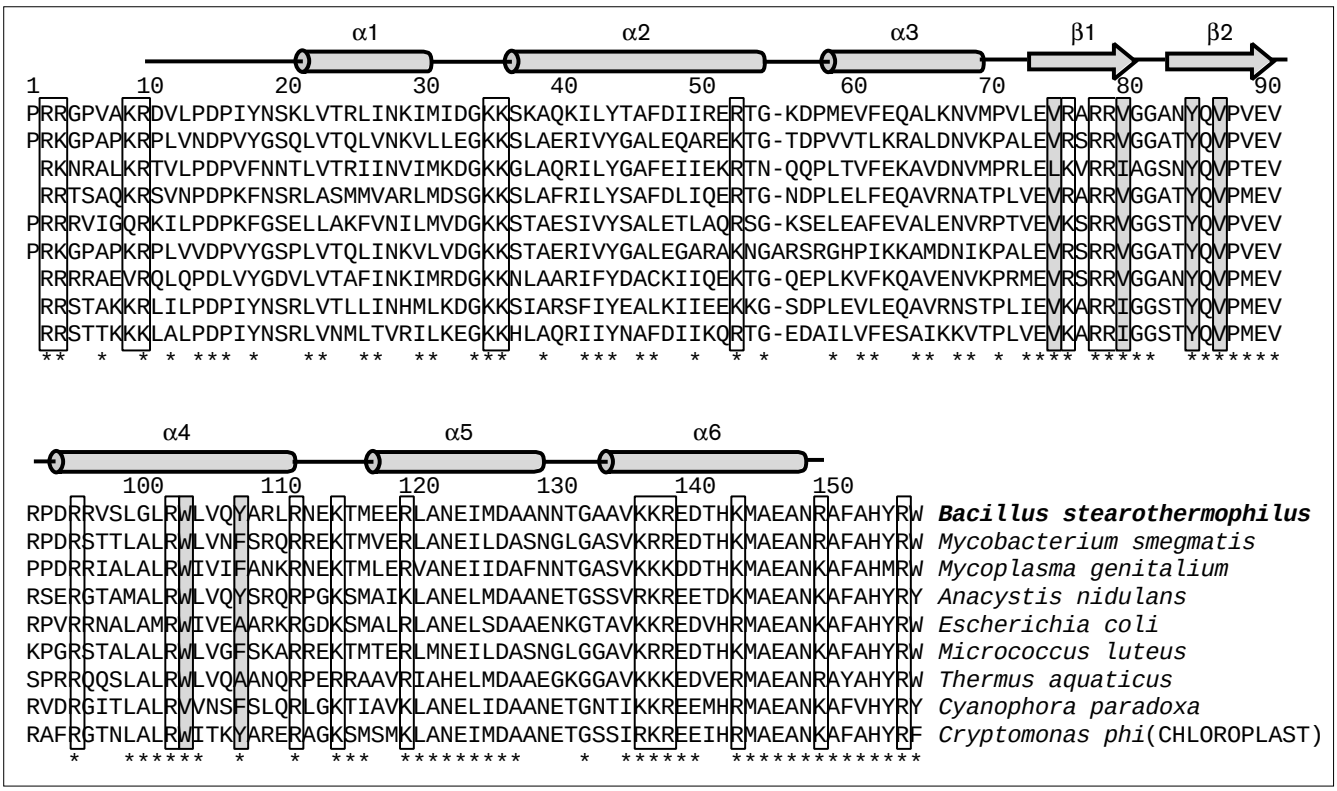
The connectivity scheme of the secondary structure units of S7 is α1–α2–α3–β1–β2–α4–α5–α6. The first two α helices, α1 (residues 20–29) and α2 (residues 35–53), the central α helix, α3 (residues 57–68), and the two α-helices following the β ribbon, α4 (residues 92–109) and α5 (residues 115–127), form a five-helix bundle structure with a pseudo-twofold symmetry axis in the centre (Figure 1). The chain directions of the first two helices (α1 and α2) and the latter two helices (α4 and α5) are reversed in this folding motif of pseudo-C_2 symmetry. Similar motifs were searched for in the protein data bank [20] using the program DALI [21], but none were detected except for a DNA architectural factor, HU and a few other proteins with limited similarities (discussed later). The central helix, α3, which runs perpendicularly to the pseudo-twofold axis, closes the open base of the five-helix bundle.

Figure 1



The overall fold of ribosomal protein S7. **(a)** Schematic stereo view of ribosomal protein S7 drawn using MOLSCRIPT [67] and Raster 3D [68]. The molecule consists of an α-helical domain (red and blue) with a β ribbon (green) extending from this domain. The α-helical domain consists of five α helices which form a five-helix bundle structure with a pseudo-twofold symmetry axis in the centre. The molecule is viewed from the pseudo-twofold symmetry axis; every tenth residue is marked. The first two helices (α1 and α2) make a helix-turn-helix (HTH) motif and helices (α4 and α5) make a 'reverse HTH' motif. These helices (shown in red) are arranged in a similar way to the body of the DNA-binding protein HU. (See section 'Comparison with a DNA architectural factor'.) **(b)** The dimeric structure of DNA-binding protein HU viewed from the molecular twofold axis, showing the similar architecture of the hydrophobic core. A pair of HTH motifs (shown in red) are entangled with each other to form a tightly packed hydrophobic core domain. In HU, the open base of the structure is closed by a pair of β strands; a highly elongated pair of β-ribbon arms protrudes downwards from the core.

Figure 2



Sequence alignment of ribosomal protein S7 from *Bacillus stearothermophilus* and its homologues. The secondary structure elements indicated are those defined by the present work using the program DSSP [69]. Conserved residues are indicated by asterisks. Conserved basic and aromatic residues, candidates for rRNA binding, are boxed or shaded, respectively. Conserved hydrophobic residues in the β-ribbon arm, which are exposed to the molecular surface are also shaded.

Between helices α3 and α4, an 18-residue stretch of polypeptide chain forms an extended β-ribbon structure, comprising β1 (residues 72–79) and β2 (residues 82–89), with a natural right-handed twist. The C-terminal helix α6 (residues 132–146) extends half way along the β ribbon. There are some hydrophobic contacts between the β ribbon and α6 which involve residues Leu72, Val74, Pro87 and Ala144. The N-terminal 19 residues preceding the α1 helix extend from the hydrophobic domain. Both termini of the polypeptide chain have no clear electron density; eight residues at both the N and C termini are completely disordered and not included in the present model.

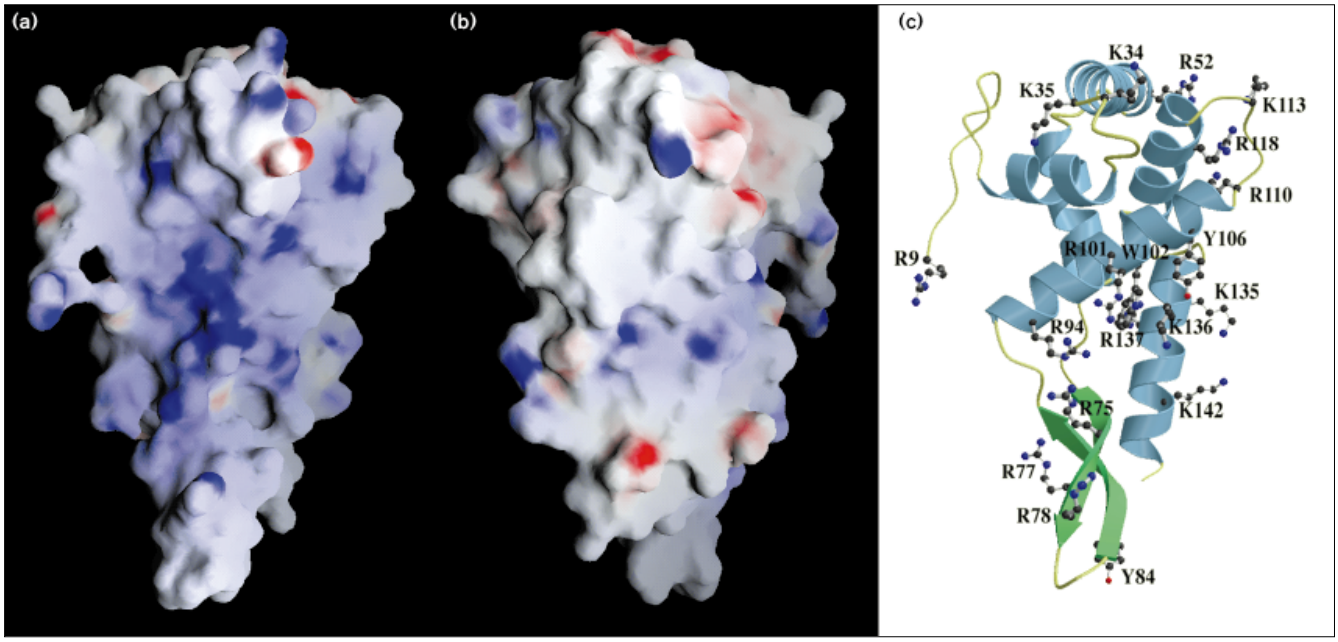
Possible functional site

The 30S subunit of the ribosome is a ribonucleoprotein complex composed of 16S rRNA and about 20 proteins. Comparative sequence analysis of 16S rRNA from several hundred organisms has led to the determination of a secondary structure model for the molecule. This model consists of three major domains, designated as 5′, central and 3′ major domains. More than 40 helix regions were also identified in this secondary structure model, which were numbered consecutively from the 5′- to the 3′-end (for example see [22]).

The 16S rRNA folds into its final structure with the help of several ribosomal proteins. S7 and S4 are two protein components responsible for initiating the folding of 16S rRNA. The former initiates the assembly at the 3′-end and the latter initiates assembly at the 5′-end, to create the head and body of the bilobal structure of the 30S subunit, respectively [11]. The binding of S7 to the 16S rRNA affects the susceptibility of nucleotides from a large region of the 3′ major domain to chemical and enzymatic probes [23], suggesting that S7 makes contacts to, and induces the folding of, rather distant portions of the primary structure of the 16S rRNA. In order to locate potential 16S rRNA-binding sites of S7, we identified patches of basic and aromatic residues within the protein that are highly conserved in the primary structures of S7 proteins from different species (Figure 2).

The highly conserved basic residues (Arg75, Arg77, Arg78, Arg94, Arg101 and Arg110) and aromatic residues (Tyr84,

Figure 3



Potential RNA-binding sites. **(a)** Molecular surface renderings of the S7 molecule. Surfaces are coloured according to their electrostatic potentials as calculated by the program GRASP [70]. The surface potential is displayed as a colour gradient from red (negative) to blue (positive), showing the relatively strong electropositive character of the putative RNA-binding site. **(b)** The view after 180° rotation from (a), showing a rather neutral surface. **(c)** Ribbon representation of the S7 molecule viewed from the same orientation as in (a). The highly conserved basic (R9, K34, K35, R52, R75, R77, R78, R94, R101, R110, R113, R118, K135, K136, R137, K142 and K148) and aromatic (Y84, W102 and Y106) residues are shown.

Trp102 and Tyr106) are clustered in the region spanning the β ribbon and α4, with their sidechains protruding from the molecular surface. These residues are localized on one face of the molecule. Calculation of the electrostatic surface potential of the S7 molecule confirmed that positively charged residues are concentrated on this one face (Figure 3). We believe that this is the main region of the S7 molecule responsible for 16S rRNA binding. This face has a concave surface, as illustrated in Figure 3, and therefore may wrap around the 16S rRNA. This putative rRNA-binding site is closely associated with the C-terminal α6 helix, in which conserved basic residues (Lys135, Lys136, Arg137 and Lys142) are also distributed. The involvement of the C-terminal region in rRNA binding has been reported in cross-linking experiments [24].

A second rRNA-binding site, also identified by the presence of conserved basic residues, is located at the N terminus of the S7 protein (Figure 2). A cross-linking experiment has shown these N-terminal residues to be involved in interactions with 16S rRNA (the rRNA is cross-linked with Lys8) [25] as well as with the antibiotic puromycin [26]. In the current electron-density map, the extreme N terminus has weak density; this region has no close contacts to the other parts of the molecule and is possibly flexible in nature.

The side of the hydrophobic core, where two loops α1–α2 and α4–α5 are situated close to each other, provides a third region for possible rRNA binding. Several conserved basic residues are clustered in this region (Lys34, Lys35 and Lys113). In addition, Met115 which has been shown to be cross-linked to rRNA by low doses of UV irradiation [24,25,27], is located at the edge of the loop α4–α5.

The orientation of S7 at the ribosome decoding centre

The most detailed morphological features of the ribosome acquired so far have been obtained by cryo-EM studies [2,3]. The cryo-EM images have been combined with the results of various experiments, such as cross-linking, mutagenesis and footprinting, to provide a detailed three-dimensional arrangement of the ribosomal RNAs [6]. How does our S7 structure correlate with this model?

Cross-linking experiments with UV irradiation or 2-iminothiolane treatment provide important information on the orientation of the S7 molecule within the ribosome. Met115 of *Escherichia coli* S7 is cross-linked to 16S rRNA upon UV irradiation [24,25,27]. The cross-linked site on the 16S rRNA is 1238–1240 [27,28]. *E. coli* S7 has also been cross-linked using 2-iminothiolane to a second site on the 16S rRNA, at nucleotides 1377–1378 [29], close to the decoding centre of the 16S rRNA. Mass spectroscopy of cross-linked

oligonucleotide–peptide complexes showed this second site to be cross-linked to Lys75 of *E. coli* S7 (Arg75 for *B. stearothermophilus*) [30]. These two cross-linking sites fix two points of S7 within the 30S ribosome model (Figure 4), while the S7 structure provides some additional constraints for the current ribosome model.

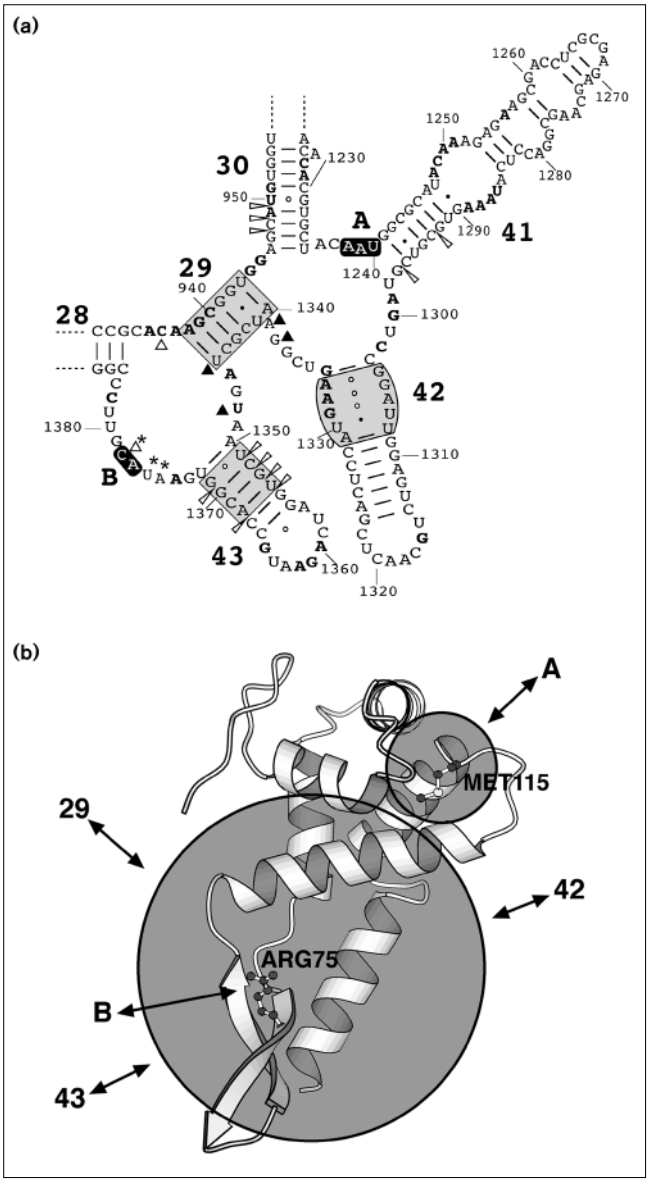
In a different set of experiments, using RNA transcripts synthesized *in vitro*, S7 has been shown to bind to a relatively wide region of the lower half of the 3′ major domain of the 16S rRNA (nucleotides 926–986 and 1219–1393) [31]. Using a series of deletion mutants, several stretches of nucleotides responsible for S7 binding were identified in this region, which spans hairpin 42, hairpin 43 and helix 29 of the rRNA (Figure 4) [32]. We propose that the main rRNA-binding site, the region spanning the β ribbon and α4, is responsible for the binding of this wide region of the 16S rRNA. This contact region is situated between the two cross-linking sites (mentioned above) in the well established secondary structure map of 16S rRNA (Figure 4). The β ribbon, possibly together with α4 and α6, grips helix 29 of the 16S rRNA, with concomitant interaction with hairpins 42 and 43, to create a tertiary structure in the 3′ major domain of 16S rRNA. The tRNA-binding sites of 16S rRNA (marked by asterisks in Figure 4) elude the grip of S7 and create part of the functional region of the ribosome.

The proposed model is consistent with the current 16S rRNA model in which the Met115 interaction site is at the 'forehead' and the Arg75 interaction site is close to the 'neck' between the head and the body of the 30S ribosomal subunit (F Mueller and R Brimacombe, personal communication). The S7 binding strongly protects helix 43 from cleavage by V1 nuclease [23]. The distal part of the 16S rRNA primary sequence (nucleotides 935–951) is also protected by S7, further suggesting that helix 29 participates in S7 binding. We attempted to fit the S7 structure into the cryo-EM model of the ribosome in order to identify interactions with 16S rRNA, the A and P site tRNAs and the mRNA. The model fitting suggested that helix 29 is, indeed, nicely bound to the concave surface of S7 with the conserved basic residues mediating the interaction (F Mueller, R Brimacombe, HH, AN, IT and SW, unpublished data). In this model, helix 43 runs along the β arm to wrap around the side of S7; the most distal part of the β arm is exposed to the tRNA sitting at the P site and the upstream region of the codon–anticodon interaction site of the mRNA.

Comparison with a DNA architectural factor

Crystallographic studies of ribosomal proteins have revealed several important aspects of the interactions between ribosomal proteins and rRNAs and have provided some information as to the evolution of this class of proteins [33,34]. Ribosomal proteins represent the relatively ancient form of

Figure 4



Correlation with biochemical experiments. **(a)** Secondary structure model of the lower half of the 16S rRNA 3′ major domain, showing the location of contact sites with S7 and tRNA. The secondary structure and helix numbers are those of Brimacombe [71]. S7 cross-link sites, as identified in the literature [28–30], are drawn in white letters (also marked by A and B). Three helices 29, 42 and 43, important for S7 binding as identified by the experiments using site-directed mutagenesis [31,32], are marked by grey shaded regions. Nucleotides that show reduced activity towards chemical probes upon S7 binding are indicated in bold letters and open arrowheads indicate the reduced reactivity sites towards V1 nuclease [23]. Nucleotides found to cross-link to tRNA [72–74] bound at the A, P and E sites are marked by small open triangles, closed triangles and asterisks, respectively. **(b)** The proposed 16S rRNA-binding sites of S7. Circles represent the binding regions. The sidechains of Met115 and Arg75, which have been identified as cross-linked residues, are shown. They are cross-linked to A and B sites of (a), respectively. The proposed main rRNA-binding site is responsible for the binding to the wide region including helices 29, 42 and 43.

protein folding. Many nucleic acid binding motifs seem to have their origins in the ribosomal proteins. The most recurring rRNA-binding motif is a split β-α-β fold, originally found in the RNA-binding domain of the spliceosomal protein U1A [35]. This motif has been found in six of the currently available ribosomal protein structures. Ribosomal proteins S17 [36] and S1 [37] consist of a five-stranded twisted β sheet, a fold similar to that of bacterial cold shock protein [38]. Helix-turn-helix (HTH) motifs have also been found in two ribosomal proteins L7/L12 and L11 [39,40]. Furthermore, amino acid sequence analysis has suggested that zinc-finger motif is likely to be found in several ribosomal proteins [41,42], although no three-dimensional structure has yet confirmed this.

While the overall folding topology of S7 is unique, it has an intriguing similarity with HU protein, a prokaryotic DNA-architectural factor. The hydrophobic core domain of HU consists of a tightly entangled pair of HTH motifs from two protomers, as shown in Figure 1 [43,44]. This topology is unique and no similar structure has been found so far. Surprisingly, the helical packing of the hydrophobic core domain of S7 is similar to that of the HU dimer. In the case of S7, the core structure is created by a single polypeptide chain, and the chain direction is reversed at the mid point as mentioned earlier. The orientation of the four helices, however, is quite similar to that of the HU dimer. The two N-terminal helices, α1 and α2, of S7 create an HTH motif, although the turn region is shorter than that observed in canonical HTH motifs [45] by two residues. At the HTH region, the amino acid sequences of the two proteins share about 30% identity (11/37). HTH motifs were originally found in three prokaryotic DNA-binding proteins [46–48], however, subsequent analyses have revealed that the motif is frequently found in a variety of proteins, including two ribosomal proteins as mentioned above. Therefore, what should be emphasized here is not the coexistence of the HTH motifs, but the similarities of the architecture where the two HTH motifs (the second 'HTH' of S7 is reversed in its chain direction) are closely associated to form a hydrophobic core domain in the two proteins.

With only limited similarities in the sequences and molecular architecture, one cannot expect to identify an evolutionary relationship between these two proteins. However, it is interesting to note that the two proteins have similar architectural functions. One of the roles of ribosomal proteins is to facilitate the folding of rRNA. This role is especially prominent for S7, which is the only primary rRNA-binding protein in the head of the small subunit [11]. The adjustment of the conformation of 16S rRNA mediated by S7 is a necessary step for the subsequent binding of other rRNA-binding proteins, such as S9 and S19 [23]. S7 also creates a site in the rRNA for tRNA binding [23]. The physiological role of the DNA-binding protein HU is also considered to be structural in that it facilitates distortions in different DNA-binding sequences in order to create an environment for other nucleoproteins to exert their functions [49,50]. Unlike most DNA-binding proteins, one member of the HU family of proteins, integration host factor (IHF), is known to bind to DNA in the minor groove [51]. IHF binds to the DNA via a pair of β ribbon arms and creates a sharp bend in the DNA with the interaction of the molecular surface of the hydrophobic core domain. In order to compensate for the energy required for bending the DNA chain, the positively charged tight body of the protein is placed on the inside of the bend, to neutralize the negatively charged phosphate backbone asymmetrically, while a hydrophobic residue on the β-ribbon arm is inserted between the stacking bases on the opposite side of the DNA.

Naturally, double stranded B-form DNA has a more regular structure than rRNA. In addition, HU protein exhibits higher symmetry than S7, as it forms dimers compared to the monomeric S7. The β-ribbon arm of S7 extends almost perpendicularly to its pseudo-twofold axis, whereas in HU a pair of arms extend along the twofold axis. Therefore, it is likely that the folding mechanism of rRNA mediated by S7 differs from that of the HU-mediated DNA folding. However, one can speculate a similar architectural role for the α-helical core domain. One of the rRNA contact sites of S7 is Met115, which is located at the turn region between α4 and α5. This residue cross-links to the oligonucleotide stretch 1238–1240 of 16S rRNA which is presumed to lie at the edge of the head of the 30S subunit, as mentioned in the previous section. A central part of the long helix 41 of 16S rRNA is protected by S7 binding (Figure 4). Assuming proximity between the 30S and 50S subunits in the vicinity of S7, it appears that this long helix folds in a direction away from the 50S subunit (F Mueller and R Brimacombe, personal communication), possibly because of the interactions with the positive surface of S7. Thus, the hydrophobic core domain created by the entangled HTH motifs may be most suited to creating the interior core which directs the bending and the folding of the nucleic acids. In this way it may function as an architectural factor.

Biological implications

Ribosomes are universal cell organelles facilitating the crucial function of protein biosynthesis. Despite the extensive studies performed since the 1950's, when ribosomes were first discovered, little is known about the detailed mechanism by which mRNAs are decoded and the corresponding proteins synthesized. In particular, many different types of experiment have been performed to delineate the region of the ribosome decoding site. One such approach used photolabelled tRNAs containing photoreactive nucleosides at the tRNA anticodon loop. The rRNAs and proteins coming into contact with the

labelled anticodon loop were subsequently marked by the probes and analyzed. These experiments revealed that ribosomal protein S7 is the protein component which is almost exclusively marked by the tRNAs bound at the aminoacyl-tRNA (A), peptidyl-tRNA (P) and exit (E) sites of the ribosome, suggesting that S7 is located in the vicinity of the ribosome decoding site [13–16].

The three-dimensional structure analysis of S7 reported here, shows that the molecule has a concave surface in which conserved basic and aromatic residues are clustered; this surface is proposed to be the 16S rRNA-binding site. The structure comprises an α-helical core linked to a β-ribbon arm and is consistent with the results of previous cross-linking experiments between 16S rRNA and S7. A preliminary model fitting attempt was performed to try and identify the interactions between S7 and 16S rRNA, tRNAs and mRNA. The model gave indications that the β-ribbon arm lies right next to the anticodon loop of the P site tRNA and the mRNA is located just upstream of the codon–anticodon base pairs. This suggests that the β ribbon of S7 might act as a gate keeper for the P site tRNA. S7 can now be used as a new atomic-resolution probe for elucidating the more detailed structure of the decoding centre. One attractive idea is the use of S7 as molecular scissors [52] which act upon bound tRNAs to ensure their displacement from the ribosome decoding site.

The architectural role of protein S7 for initiating the assembly of the head of the 30S ribosomal subunit has also been well established. The folding of 16S rRNA mediated by S7 binding is the necessary step for other proteins to bind and to create a functional 30S particle [11,12]. The structure of S7 contains features which may explain how this molecule binds to 16S rRNA and initiates its folding to create the head of the 30S subunit. The hydrophobic core structure observed in S7 consists of two entangled helix-turn-helix motifs and the fold shows similarity to a DNA architectural factor. Thus, these two types of proteins seem to employ a similar strategy for creating the interior core which directs the bending and the folding of nucleic acids.

Materials and methods

Preparation of the selenomethionyl S7 crystal

S7 from *B. stearothermophilus* consists of 155 amino acid residues. The protein was overexpressed in *E. coli* cells, purified and crystallized as will be reported elsewhere [53]. In order to determine the crystal structure using the MAD method [54], we prepared selenomethionyl (Se-Met) substituted protein. Crystals of Se-Met S7 were obtained under similar conditions to those used for native S7. The best crystals were grown from a solution containing 0.1 M Na Hepes buffer (pH 8.2) with 4% (v/v) of 2-propanol and 2.0 M ammonium sulphate. The typical size of the crystals was about 0.3 × 0.25 × 0.25 mm^3. Cell dimensions were determined from oscillation photographs using the HKL package [55]. Although the Se-Met S7 crystals were obtained under the similar conditions to the native ones, they are not isomorphous. The Se-Met S7 crystal belongs to the space group $P2_12_12$ with cell dimensions of a = 54.40, b = 131.02 and c = 29.15 Å, whereas the space group and cell dimensions of native crystal are $P2_1$ and a = 30.55, b = 117.6, c = 56.09 Å and β = 96.2°. Assuming that the crystal contains one monomer in an asymmetric unit, the V_M value is calculated as 2.90 $Å^3$/Da, corresponding to a solvent content of 57.3% [56].

Data collection

Data collection statistics are given in Table 1. All diffraction data sets were collected from a single crystal on the MAD beamline, BM14 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Four different wavelengths (0.90007 Å [e0], 0.97881 Å [e1], 0.97906 Å [e2] and 0.97981 Å [e3]) were chosen for MAD data collection. The wavelengths were optimized to obtain a larger anomalous signal based on the fluorescence spectrum. All diffraction data were recorded on a Princeton CCD detector coupled to an image intensifier and processed with the HKL package.

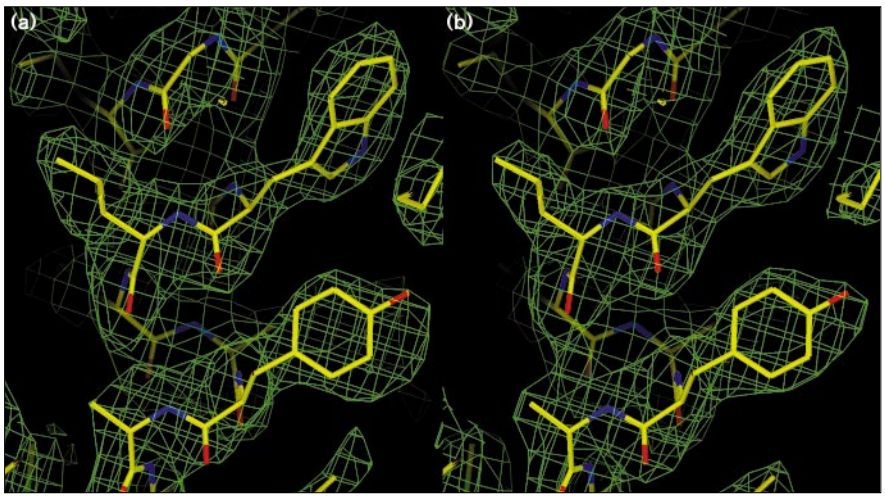
Table 1

Crystallographic data.

Data	e0	e1	e2	e3
Wavelength (Å)	0.90007	0.97881	0.97906	0.97981
Resolution (Å)	30.0–2.5	30.0–2.5	30.0–2.5	50.0–2.5
R_{merge}*	0.048 (0.138)	0.041 (0.171)	0.039 (0.176)	0.050 (0.431)
Observed reflections	53,236	45,531	45,666	42,500
Independent reflections	7,558	7,427	7,424	7,529
Completeness (%)	97.6 (92.9)	96.6 (88.2)	96.5 (87.3)	95.9 (85.5)
Multiplicity	2.8 (2.2)	2.5 (1.9)	2.5 (1.9)	2.3 (1.6)
R_{lambda}†		0.064	0.098	0.098
f'‡	−1.6	−6.6	−10.8	−4.7
f''§	3.3	5.1	5.5	0.7

Values within parentheses are for the highest resolution shell (2.54–2.50 Å). *$R_{merge} = \Sigma\Sigma_j | <I(h)> - I(h)_j | / \Sigma\Sigma_j <I(h)>$, where <I(h)> is the mean intensity of symmetry-equivalent reflections. Friedel pairs were merged as individual data. †$R_{lambda} = \Sigma\Sigma_j | |F_{\lambda j}| - |F_{\lambda 0}| | / \Sigma\Sigma_j |F_{\lambda 0}|$, where $F_{\lambda j}$ is the structure factor of the data collected at λ_j and $F_{\lambda 0}$ is the structure factor of the data collected at e0 (0.90007 Å). ‡The real part of the anomalous scattering factor of the selenium atom refined by SHARP [61]. §Imaginary part of the anomalous scattering factor of the selenium atom refined by SHARP.

Figure 5



Part of the 2.5 Å resolution electron-density maps using **(a)** MLPHARE [59] and DM [60] phases and **(b)** SHARP [61] and SOLOMON [62] phases. The refined model of the S7 coordinates is superimposed; atoms are shown in standard colours. The contour level is 1.2 times the root mean square density of the map.

Phasing

Out of six independent Se atoms, five were located from the Bijvoet anomalous difference Patterson map of the e2 data using the real-space Patterson search program RSPS [57]. Their positions were initially refined by the vector-space refinement method using the program VECREF [58]. Heavy atom parameter refinement and phase calculation were carried out using the program MLPHARE [59] by treating the data as a special case of multiple isomorphous replacement. The initial electron-density map was improved by solvent flattening and histogram mapping using the program DM [60]. The improved map was of sufficient quality to build an atomic model. As an alternative approach to heavy-atom refinement and phasing, the maximum-likelihood method was also applied to the same data sets using the program SHARP [61]. The electron-density map obtained after phase improvement by the program SOLOMON [62] gave slightly better defined details (Figure 5). The atomic model was first built using the electron-density map from MLPHARE, but at the later stages the map from SHARP was used. The results of phasing calculations are summarized in Table 2.

Model building and refinement

The atomic model was built using the graphics program O [63]. Out of the 155 residues, 135 were built into the initial electron density map. The model was refined against the remote data (e0) using the program X-PLOR [64]. The R factor for the data between 8 Å to 2.5 Å resolution dropped from 0.436 to 0.288 after the first cycle of simulated annealing and B factor refinement.

Table 2

Phasing statistics.

Data	e0		e1		e2		e3	
	isomorphous	anomalous	isomorphous	anomalous	isomorphous	anomalous	isomorphous	anomalous
MLPHARE* + DM								
R_{Cullis}†								
(centric)			0.74		0.63		0.83	
(acentric)			0.95	0.62	0.72	0.56	0.92	0.98
Phasing power‡								
(centric)				0.95		1.19		0.55
(acentric)				1.23		1.57		0.75
Figure of merit§	0.836							
SHARP + SOLOMON								
R_{Cullis}†								
(centric)			0.544		0.506		0.887	
(acentric)		0.654	0.569	0.493	0.506	0.497	0.937	0.961
Phasing power‡								
(centric)			0.240		0.213		0.866	
(acentric)		3.244	3.078	3.903	3.36	3.76	1.15	0.88
Figure of merit§	0.901							

*MAD phasing was calculated as a special case of MIRAS, where e0 data was used as native and other wavelength data were used as individual derivatives. †$R_{Cullis} = \Sigma_j | E | / \Sigma_j | |F_\lambda| - |F_{\lambda 0}| |$, where E is the lack of closure. ‡Phasing power $= < | F_H(calc) | / | E | >$, where $F_H(calc)$ is the calculated anomalous difference and E is the lack of closure. §Figure of merit after density modification.

At the current stage of refinement, the model has an R factor of 21.6% for 90% of the data between 8 Å and 2.5 Å, including 139 residues (9–147) and 25 water molecules, a total of 1145 atoms. The free R factor [65] for the remaining 10% of the data within this resolution range is 27.0%. The Ramachandran plot of the model shows that 94.5% of the residues lie within the most favoured region with no residues in disallowed regions of the plot. The root mean square (rms) deviations from standard values [66] of bond lengths and angles are 0.020 Å and 2.022°, respectively.

Accession numbers

Coordinates have been deposited in the Brookhaven Protein Data Bank with the accession code 1HUS.

Acknowledgements

We thank V Stojanoff of the ESRF and A Thompson of EMBL Grenoble Outstation for their assistance during the MAD data collection on BM14 at the ESRF. We also thank AG Murzin, PR Evans and R Brimacombe for their careful reading of the manuscript and valuable comments. This work was supported by the Proposal-Based Advanced Industrial Technology R&D Program of the New Energy and Industrial Technology Development Organization (NEDO) of Japan and by a Grant-in-Aid for the Ministry of Education, Science and Culture of Japan. AN and IT are members of the TARA project of Tsukuba University, Japan.

References

1. Gornicki, P., Nurse, K., Hellmann, W., Boublik, M. & Ofengand, J. (1984). High resolution localization of the tRNA anticodon interaction site on the *Escherichia coli* 30 S ribosomal subunit. *J. Biol. Chem.* **259**, 10493–10498.
2. Stark, H., *et al.*, & van Heel, M. (1995). The 70S *Escherichia coli* ribosome at 23 Å resolution: fitting the ribosomal RNA. *Structure* **3**, 815–821.
3. Frank, J., *et al.*, & Agrawal, R.K. (1995). A model of protein synthesis based on cryo-electron microscopy of the *E. coli* ribosome. *Nature* **376**, 441–444.
4. Agrawal, R.K., *et al.*, & Frank, J. (1996). Direct visualization of A-, P-, and E-site transfer RNAs in the *Escherichia coli* ribosome. *Science* **271**, 1000–1002.
5. Stark, H., *et al.*, & van Heel, M. (1997). Arrangement of tRNAs in pre- and post-translocational ribosomes revealed by electron cryomicroscopy. *Cell* **88**, 19–28.
6. Brimacombe, R. (1995). The structure of ribosomal RNA: a three-dimensional jigsaw puzzle. *Eur. J. Biochem.* **230**, 365–383.
7. Capel, M.S., *et al.*, & Sillers, I.Y. (1987). A complete mapping of the proteins in the small ribosomal subunit of *Escherichia coli*. *Science* **238**, 1403–1406.
8. Brimacombe, R., Atmadja, J., Stiege, W. & Schüler, D. (1988). A detailed model of the three-dimensional structure of *Escherichia coli* 16 S ribosomal RNA *in situ* in the 30 S subunit. *J. Mol. Biol.* **199**, 115–136.
9. Schüler, D. & Brimacombe, R. (1988). The *Escherichia coli* 30S ribosomal subunit; an optimized three-dimensional fit between the ribosomal proteins and the 16S RNA. *EMBO J.* **7**, 1509–1513.
10. Pondkowinski, J. & Gornicki, P. (1989). Ribosomal proteins S7 and L1 are located close to the decoding site of *E. coli* ribosome – affinity labeling studies with modified tRNAs carrying photoreactive probes attached adjacent to the 3′-end of the anticodon. *Nucleic Acids Res.* **17**, 8767–8782.
11. Nowotny, V. & Nierhaus, K.H. (1988). Assembly of the 30S subunit from *Escherichia coli* ribosomes occurs via two assembly domains which are initiated by S4 and S7. *Biochemistry* **27**, 7051–7055.
12. Mandiyan, V., Tumminia, S., Wall, J.S., Hainfeld, J.F. & Boublik, M. (1989). Protein-induced conformational changes in 16 S ribosomal RNA during the initial assembly steps of the *Escherichia coli* 30 S ribosomal subunit. *J. Mol. Biol.* **210**, 323–336.
13. Abdurashidova, G.G., Tsvetkova, E.A. & Budowsky, E.I. (1989). Nucleotide residues of tRNA, directly interacting with proteins within the complex of the 30 S subunit of *E. coli* ribosome with poly(U) and NAcPhe-tRNA^{Phe}. *FEBS Lett.* **243**, 299–302.
14. Abdurashidova, G.G., Tsvetkova, E.A. & Budowsky, E.I. (1990). Determination of tRNA nucleotide residues directly interacting with proteins in the post- and pretranslocated ribosomal complexes. *FEBS Lett.* **269**, 398–401.
15. Sylvers, L.A., Kopylov, A.M., Wower, J., Hixson, S.S. & Zimmermann, R.A. (1992). Photochemical cross-linking of the anticodon loop of yeast tRNA^{Phe} to 30S-subunit protein S7 at the ribosomal A and P sites. *Biochimie* **74**, 381–389.
16. Döring, T., Mitchell, P., Osswald, M., Bochkariov, D. & Brimacombe, R. (1994). The decoding region of 16S RNA; a cross-linking study of the ribosomal A,P and E sites using tRNA derivatized at position 32 in the anticodon loop. *EMBO J.* **13**, 2677–2685.
17. Stade, K., Rinke-Appel, J. & Brimacombe, R. (1989). Site-directed cross-linking of mRNA analogues to the *Escherichia coli* ribosome; identification of 30S ribosomal components that can be cross-linked to the mRNA at various points 5′ with respect to the decoding site. *Nucleic Acids Res.* **17**, 9889–9908.
18. Dean, D., Yates, J.L. & Nomura, M. (1981). Identification of ribosomal protein S7 as a repressor of translation within the *str* operon of *E. coli*. *Cell* **24**, 413–419.
19. Saito, K. & Nomura, M. (1994). Post-transcriptional regulation of the *str* operon in *Escherichia coli*. Structural and mutational analysis of the target site for translational repressor S7. *J. Mol. Biol.* **235**, 125–139.
20. Bernstein, F.C. (1978). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
21. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
22. Moore, P.B. (1988). The ribosome returns. *Nature* **331**, 223–227.
23. Powers, T., Changchien, L.M., Craven, G.R. & Noller, H.F. (1988). Probing the assembly of the 3′ major domain of 16 S ribosomal RNA. Quaternary interactions involving ribosomal proteins S7, S9 and S19. *J. Mol. Biol.* **200**, 309–319.
24. Ehresmann, B., Reinbolt, J., Backendorf, C., Tritsch, D. & Ebel, J. (1976). Studies of the binding sites of *Escherichia coli* ribosomal protein S7 with 16S RNA by ultraviolet irradiation. *FEBS Lett.* **67**, 316–319.
25. Urlaub, H., Kruft, V., Bischof, O., Müller, E.C. & Wittmann-Liebold, B. (1995). Protein-rRNA binding features and their structural and functional implications in ribosomes as determined by cross-linking studies. *EMBO J.* **14**, 4578–4588.
26. Bischof, O., Kruft, V. & Wittmann-Liebold, B. (1994). Analysis of the puromycin binding site in the 70 S ribosome of *Escherichia coli* at the peptide level. *J. Biol. Chem.* **269**, 18315–18319.
27. Möller, K., Zwieb, C. & Brimacombe, R. (1978). Identification of the oligonucleotide and oligopeptide involved in an RNA–protein crosslink induced by ultraviolet irradiation of *Escherichia coli* 30S ribosomal subunits. *J. Mol. Biol.* **126**, 489–506.
28. Zwieb, C. & Brimacombe, R. (1979). RNA–protein cross-linking in *Escherichia coli* 30S ribosomal subunits: precise localization of the nucleotide in 16S RNA which is coupled to protein S7 by ultraviolet irradiation. *Nucleic Acids Res.* **6**, 1775–1790.
29. Wower, I. & Brimacombe, R. (1983). The localization of multiple sites on 16S RNA which are cross-linked to proteins S7 and S8 in *Escherichia coli* 30S ribosomal subunits by treatment with 2-iminothiolane. *Nucleic Acids Res.* **11**, 1419–1437.
30. Urlaub, H., Thiede, B., Müller, E.C., Brimacombe, R. & Wittmann-Liebold, B. (1997). Identification and sequence analysis of contact sites between ribosomal proteins and rRNA in *Escherichia coli* 30S subunits by a new approach using matrix-assisted laser desorption/ionization-mass spectrometry combined with N-terminal microsequencing. *J. Biol. Chem.* **272**, 14547–14555.
31. Dragon, F. & Brakier-Gingras, L. (1993). Interaction of *Escherichia coli* ribosomal protein S7 with 16S rRNA. *Nucleic Acids Res.* **21**, 1199–1203.
32. Dragon, F., Payant, C. & Brakier-Gingras, L. (1994). Mutational and structural analysis of the RNA binding site for *Escherichia coli* ribosomal protein S7. *J. Mol. Biol.* **244**, 74–85.
33. Davies, C., Ramakrishnan, V. & White, S.W. (1996). Structural evidence for specific S8–RNA and S8–protein interactions within the 30S ribosomal subunit: ribosomal protein S8 from *Bacillus stearothermophilus* at 1.9 Å resolution. *Structure* **4**, 1093–1104.
34. Liljas, A. & Garber, M. (1995). Ribosomal proteins and elongation factors. *Curr. Opin. Struct. Biol.* **5**, 721–727.
35. Nagai, K., Oubridge, C., Jessen, T.H., Li, J. & Evans, P.R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* **348**, 515–520.
36. Jaishree, T.N., Ramakrishnan, V. & White, S.W. (1996). Solution structure of prokaryotic ribosomal protein S17 by high-resolution NMR spectroscopy. *Biochemistry* **35**, 2845–2853.

37. Bycroft, M., Hubbard, T.J.P., Proctor, M., Freund, S.M.V. & Murzin, A.G. (1997). The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88**, 235–242.
38. Schindelin, H., Marahiel, M.A. & Heinemann, U. (1993). Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. *Nature* **364**, 164–168.
39. Leijonmarck, M. & Liljas, A. (1987). Structure of the C-terminal domain of the ribosomal protein L7/L12 from *Escherichia coli* at 1.7 Å. *J Mol. Biol*. **195**, 555–579.
40. Markus, M.A., Hinck, A.P., Huang, S., Draper, D.E. & Torchia, D.A. (1997). High resolution solution structure of ribosomal protein L11-C76, a helical protein with a flexible loop that becomes structured upon binding to RNA. *Nat. Struct. Biol*. **4**, 70–77.
41. Chan, Y.-L., Suzuki, K., Olvera, J. & Wool, I.G. (1993). Zinc-finger-like motifs in rat ribosomal proteins S27 and S29. *Nucleic Acids Res*. **21**, 649–655.
42. Herfurth, E., Briesemeister, U. & Wittmann-Liebold, B. (1994). Complete amino acid sequence of ribosomal protein S14 from *Bacillus stearothermophilus* and homology studies to other ribosomal proteins. *FEBS Lett*. **351**, 114–118.
43. Tanaka, I., Appelt, K., Dijk, J., White, S.W. & Wilson, K.S. (1984). 3 Å Resolution structure of a protein with histone-like properties in prokaryotes. *Nature* **310**, 376–381.
44. Vis, H., Mariani, M., Vorgias, C.E., Wilson, K.S., Kaptein, R. & Boelens, R. (1995). Solution structure of the HU protein from *Bacillus stearothermophilus*. *J. Mol. Biol*. **254**, 692–703.
45. Harrison, S.C. & Aggarwal, A.K. (1990). DNA recognition by proteins with the helix-turn-helix motif. *Annu. Rev. Biochem*. **59**, 993–969.
46. McKay, D.B. & Steitz, T.A. (1981). Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* **290**, 744–749.
47. Anderson, W.F., Ohlendorf, D.H., Takeda, Y. & Matthews, B.W. (1981). Structure of the cro repressor from bacteriophage λ and its interaction with DNA. *Nature,* **290**, 754–758.
48. Pabo, C.O. & Lewis, M. (1982). The operator-binding domain of λ repressor: structure and DNA recognition. *Nature* **298**, 443–447.
49. Flashner, Y. & Gralla, J.D. (1988). DNA dynamic flexibility and protein recognition: differential stimulation by bacterial histone-like protein HU. *Cell* **54**, 713–721.
50. Johnson, R.C., Bruist, M.F. & Simon, M.I. (1986). Host protein requirements for *in vitro* site-specific DNA inversion. *Cell* **46**, 531–539.
51. Rice, P.A., Yang, S., Mizuuchi, K. & Nash, H.A. (1996). Crystal structure of an IHF–DNA complex: a protein-induced DNA U-turn. *Cell* **87**, 1295–1306.
52. Heilek, G.M. & Noller, H.F. (1996). Site-directed hydroxyl radical probing of the rRNA neighbourhood of ribosomal protein S5. *Science* **272**, 1659–1662.
53. Harada, N., Sano, K., Kimura, M., Hosaka, H., Nakagawa, A. & Tanaka, I. (1997). Crystallization and preliminary X-ray crystallographic study of the ribosomal protein S7 from *Bacillus stearothermophilus*. *J. Struct. Biol*., in press.
54. Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58.
55. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillations mode. *Methods Enzymol.* **276**, 307–326.
56. Matthews, B.W. (1968). Solvent contents of protein crystals. *J. Mol. Biol*. **33**, 491–497.
57. Knight, S. (1989). *Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase – A Structural Study.* Thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden.
58. Tickle, I. (1991). Refinement of single isomorphous replacement heavy-atom parameters in Patterson vs reciprocal space. In *Isomorphous Replacement and Anomalous Scattering. Proceedings of the CCP4 Study Weekend, 25–26. January 1991.* (Wolf, W., Evans, P.R. & Leslie, A.G.W., eds), pp. 87–95, SERC Daresbury Laboratory, Warrington, UK.
59. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy-atom parameters. In *Isomorphous Replacement and Anomalous Scattering. Proceedings of the CCP4 Study Weekend, 25–26. January 1991.* (Wolf, W., Evans, P.R. & Leslie, A.G.W., eds), pp. 80–86, SERC Daresbury Laboratory, Warrington, UK.
60. Cowtan, K. & Main, P. (1996). Phase combination and cross validation in iterated density-modification calculations. *Acra Cryst. D* **52**, 43–48.
61. de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. *Methods Enzymol.* 276, 472–494.
62. Abrahams, J.P. & Leslie, A.G.W. (1996). Methods used in the structure determination of bovine mitochondrial F_1 ATPase. *Acta Cryst. D* **52**, 30–42.
63. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 489–501
64. Brünger, A.T. (1993). *X-PLOR Version 3.1 A system for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT, USA.
65. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475.
66. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A* **47**, 392–400.
67. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.
68. Merrit, E.A. & Murphy, M.E.P. (1994). Raster3D version 2.0, a program for photorealistic molecular graphics. *Acta Cryst. D* **50**, 869–873.
69. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
70. Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.
71. Brimacombe, R. (1991). RNA–protein interactions in the *E .coli* ribosome. *Biochimie* **73**, 927–936.
72. Rinke-Appel, J., Juenke, N., Osswald, M. & Brimacombe, R. (1995). The ribosomal environment of tRNA: Crosslinks to rRNA from positions 8 and 20:1 in the central fold of tRNA located at the A, P, or E site. *RNA* **1**, 1018–1028.
73. Osswald, M., Döring, T. & Brimacombe, R. (1995). The ribosomal neighbourhood of the central fold of tRNA: cross-links from position 47 of tRNA located at the A, P or E site. *Nucleic Acids Res.* **23**, 4635–4641.
74. Mueller, F., *et al*., & Brimacombe, R. (1995). Getting closer to an understanding of the three-dimensional structure of ribosomal RNA. *Biochem. Cell Biol*. **73**, 767–773.